Mechanisms of song perception in oscine birds

Daniel P. Knudsen [a], Timothy Q. Gentner [b],*

[a] Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA 92093, United States
[b] Department of Psychology, University of California, San Diego, La Jolla, CA 92093, United States

ARTICLE INFO

ABSTRACT

Songbirds share a number of parallels with humans that make them an attractive model system for studying the behavioral and neurobiological mechanisms that underlie the learning and processing of vocal communication signals. Here we review the perceptual and cognitive mechanisms of audition in birds, and emphasize the behavioral and neural basis of song recognition. Where appropriate, we point out a number of intersections with human vocal communication behavior that suggest common mechanisms amenable to further study, and limitations of birdsong as a model for human language.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Like other communication signals, one adaptive function of birdsong is to influence the behavior of others, usually conspecific individuals (Kroodsma & Miller, 1996). Communication signals achieve this function by transmitting information between the sender of the signal and the receivers. The success of this transmission rests on predictability. When a sender produces a specific signal, in this case a song, it does so under the expectation that the signal will elicit a predictable (i.e. intended) behavior in the receiver. Without the predictable correspondence between production and perception, signals would lose their functionality. Thus, the presence of a functional signal implies a reliable correspondence between production and perception mechanisms shaped and maintained by selection pressures. This correspondence confers a special status to communication signals. Like other natural stimuli, communication signals are often physically complex. Unlike most complex natural stimuli, however, many of the physical dimensions along which communication signals vary can be directly tied to adaptive behaviors. Research in oscine birds has capitalized on this idea to study the mechanisms underlying the perception and cognition of complex natural stimuli (song) in the context of natural behaviors.

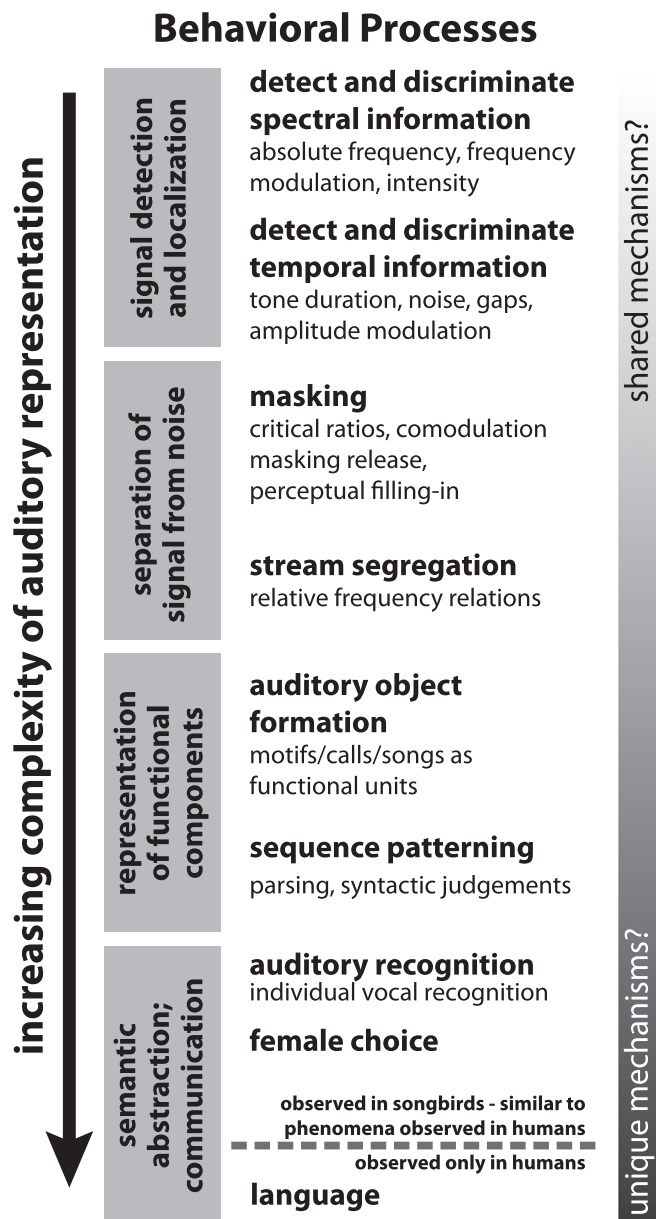What follows is an account of our current understanding about the ways in which the songbird auditory system interprets a con-tinuous stream of acoustic information as a collection of behaviorally relevant communication signals and uses this information to affect behavior (Fig. 1). Parallels will be drawn to human vocal communication, and we will present a case for the use of songbirds as a model of certain aspects of human language processing. We first provide a broad overview of perceptual psychophysics in songbirds so that one can appreciate the strong perceptual similarities between birds and humans. We then review the behavioral and neurophysiological work on conspecific song perception, focusing on individual vocal recognition mechanisms in European starlings, a species of songbird.

2. Perceptual psychoacoustics

It is helpful for any description of a complex system to begin with a characterization of general abilities along fundamental dimensions. Common dimensions along which acoustic signals are deconstructed are frequency, amplitude, and time, and it is instructive to understand sensory processing at this level. The goal of such research is to inform our understanding of how more 'atomic' descriptions of sounds give rise to the perception of more complex behaviorally relevant features. Such studies provide an important context for studies of more complex signals, and establish the range over which any perceptual ability can operate. The following section provides a brief description of avian psycho-acoustics studies supporting the notion that humans and birds experience a similar acoustic world. More thorough reviews are available (Dooling, 1982, 1992, 2000; Fay, 1988).

* Corresponding author. Address: UCSD, Department of Psychology, 9500 Gilman Dr. MC0109, La Jolla, CA 92093-0109, United States.
E-mail address: tgentner@ucsd.edu (T.Q. Gentner).

Behavioral Processes

detect and discriminate spectral information
absolute frequency, frequency modulation, intensity

detect and discriminate temporal information
tone duration, noise, gaps, amplitude modulation

masking
critical ratios, comodulation masking release, perceptual filling-in

stream segregation
relative frequency relations

auditory object formation
motifs/calls/songs as functional units

sequence patterning
parsing, syntactic judgements

auditory recognition
individual vocal recognition

female choice

observed in songbirds - similar to phenomena observed in humans
- - - - - - - - - - - - - -
observed only in humans

language

signal detection and localization

separation of signal from noise

representation of functional components

semantic abstraction; communication

shared mechanisms?

unique mechanisms?

increasing complexity of auditory representation

Fig. 1. Functional auditory behaviors in songbirds. The gray boxes represent the functions that the auditory system must perform to mediate behavior. Text to the right of these functions gives examples of phenomena observed in songbirds (see text). Although this review focuses on songbird auditory processing, many of the functions of the auditory system are likely conserved across a range of vertebrates, including humans. As behavioral complexity increases, so does the likelihood that particular mechanisms are unique to different species. The similarities and the differences between species yield powerful comparative hypotheses about the behavioral and neural mechanisms for auditory perception and cognition in vertebrates.

2.1. Spectral sensitivity

Audibility curves, which describe the loudness (sound pressure level, SPL) required for detectability as a function of frequency, have the same basic shape in many songbird species (Dooling, Okanoya, Downing, & Hulse, 1986; Okanoya & Dooling, 1987). The threshold sensitivity for pure tones in most songbird species is best around 2–5 kHz, increases gradually as the frequency of the tone becomes lower, and increases quite sharply as the frequency of the tone rises. The typical high-frequency cutoff for songbirds is 8–11 kHz. Although audibility thresholds are lower in humans than in birds, the overall shape of the audibility curve is similar.

In addition to hearing single tones, the auditory system must also discriminate between different tones. Understanding the minimal detectable differences between tones can point to the types of frequency modulations within a signal that are available for a species to use in vocal communication. In general, birds are quite sensitive to changes in frequency, and can discriminate a change in frequency as small as 1% (Dooling, 1982), while humans have even lower detection thresholds across the range of audible frequencies. Birds' sensitivity to frequency changes also depends on the type of frequency modulation and range of carrier frequencies (Langemann & Klump, 1992), results again similar to observations made in human studies (Demany & Semal, 1989; Fastl, 1978). To detect a change in intensity (loudness) between two successive tones, birds require a difference of about 3 dB, humans about 1 dB (Dooling, 1982). While humans have quantitatively lower detection thresholds for frequency and loudness discrimination, the similarity of findings in humans and birds point to auditory systems that are qualitatively similar in the range of psychophysically observable spectral sensitivities.

2.2. Temporal sensitivity

Most acoustic signals unfold over time, and processing in the temporal domain is therefore particularly important. Temporal processing abilities of songbirds have been studied in a variety of ways. A simple measure is the detectability of a sound as its temporal duration increases. In general, the longer the tone is played for, the lower the SPL needed for detection. Consistent with findings from a host of other animals including humans (see Brown & Maloney, 1986), birds' thresholds for hearing a pure tone improve as the duration is increased from a few milliseconds to 200–300 ms (Dooling, 1980). Another common measure of temporal acuity, known as the gap detection threshold, measures the minimum temporal interval that can be detected between two sounds. Several studies from birds show gap detection thresholds ranging from 2 to 3 ms, which is similar to thresholds found in humans (see Klump & Maier, 1989). This suggests that intervals in natural vocalizations less than 2–3 ms may not be perceived. Duration discrimination measures, which describe an organism's ability to determine whether one sound has a longer duration than another, are also similar between birds and humans (Maier & Klump, 1990).

2.3. Masking

In psychoacoustics, the critical ratio describes the ability of an organism to perceive a tone in a noisy background. It is defined as the SPL of a target tone needed for detection divided by the SPL of the background masking noise. Critical ratios are a function of the frequency of the target tone, and in accordance with previous studies in humans and in other mammals, critical ratios in most songbirds increase at about 3 dB per octave (Dooling et al., 1986; Langemann, Klump, & Dooling, 1995; Okanoya & Dooling, 1987). Most songbirds' critical ratio curves show a similar shape to those of humans and other mammals, though humans show lowered threshold levels on the order of a few decibels (Okanoya & Dooling, 1987).

Comodulation masking release (CMR) is a slightly more complex masking phenomenon that has been described in both birds and humans. CMR occurs when sounds that are modulated together across time serve to release each other from masking by overlapping noise. CMR is measured as the effective decrease in threshold SPL afforded by the comodulation. CMR has been proposed as mechanism for auditory stream segregation (discussed below), as sounds that are produced from the same source take the same path to the listener and are therefore modulated

identically. Starlings show robust CMR, and in many ways are similar to humans in terms of amount of masking release and the frequency and bandwidth dependencies of the release (Klump & Langemann, 1995).

3. Auditory cognition in birds

Within the broad framework of acoustic processes laid out by the results of psychoacoustic studies, birds (and any other organism that uses acoustic communication) face a number of specific challenges. The auditory system must first detect a potentially relevant signal against backgrounds of environmental noise or other potentially relevant signals. Signals must be processed to determine their relevance in specific behavioral contexts, and mapped onto appropriate behaviors. These behaviors are often mediated by learning, memory and attentional mechanisms, implying a suite of cognitive processes that act on complex stimulus representations such as stream segregation, discrimination, recognition, categorization, temporal pattern detection, and decision processes. The acoustic diversity of birdsong makes the system attractive for studies of the behavioral and neurobiological bases for many of these cognitive processes.

3.1. Stream segregation

One challenge to the auditory system is the separation of incoming auditory information into behaviorally relevant groups, or streams. This is referred to as stream segregation, and under most natural conditions equates to separation of streams based on originating source. In order to effectively use the acoustic information in one's environment, it is helpful to group acoustic information arising from the same source, and to treat this information as separate from information arising from other sources. In human language, for example, this ability allows one to follow the vocalizations of a speaker across time in a noisy environment.

Stream segregation is one of the main tasks of any perceptual system engaged in vocal communication, and it is therefore no surprise that songbirds are capable of separating one auditory stream from another in a variety of conditions. Like humans (Bregman, 1990), starlings can hear competing streams emerge from a sequence of alternating tones when the frequency difference between putative "streams" is large (MacDougall-Shackleton, Hulse, Gentner, & White, 1998). Likewise, starlings continue to recognize long segments of conspecific songs when noise is added in the form of other starling songs, the songs of other species, or an entire 'dawn chorus' (Hulse, MacDougall-Shackleton, & Wisniewski, 1997). Canaries can maintain accurate recognition of conspecific song in the face of very loud broadband white noise (Appellants, Gentner, Hulse, Balthazart, & Ball, 2005). The neural correlates of the so-called cocktail party effect, that is the extraction of a meaningful communication signal from noise, have been investigated in starlings and zebra finches (Bee & Klump, 2004; Narayan et al., 2007).

3.2. Song discrimination and recognition

Many species of bird are adept at discriminating songs of their own species as well as songs from other species. Within this context, birds are generally more adept at discriminating among different conspecific songs than among heterospecific songs (Cynx & Nottebohm, 1992; Dooling, Brown, Klump, & Okanoya, 1992; Sinnott, Sachs, & Hienz, 1980). Fieldwork in sparrows is consistent with these laboratory findings, and suggests that different species may rely, at least for species recognition, on those acoustic features that most reliably distinguish their own songs from those of other species in the same sound environment (Nelson & Marler, 1990). For example, frequency appears to be the most important cue for species recognition among field sparrows (Nelson, 1989). Recognition systems appear to be tuned to the particular acoustic space of one's own species' vocalizations (Nelson, 1988), but also maintain more general capacities to distinguish among many different types of song. These biases for the discrimination of species-specific vocalizations, and hence mechanisms for vocal recognition, likely result from evolutionary or ontogenic changes in the (presumably) central processing mechanisms that underlie higher-level auditory perception. This is consistent with the more general assumption that the perceptual and cognitive processes underlying individual vocal recognition, a widespread phenomenon among songbirds (reviewed by Stoddard (1996)), operate on complex (multi-dimensional) acoustic signals. Within the same species many acoustic features may be relevant for classification of song components (Sturdy, Phillmore, & Weisman, 2000). Recent laboratory studies of European starlings have addressed these questions directly by determining more precisely the form of the acoustic signal controlling recognition in this species.

3.2.1. Individual vocal recognition in starlings

Male starlings present their songs in long episodes of continuous singing referred to as bouts. Song bouts, in turn, are composed of much smaller acoustic units referred to as motifs (Adret-Hausberger & Jenkins, 1988; Eens, Pinxten, & Verheyen, 1991), which in turn are composed of still smaller units called notes. Notes can be broadly classified by the presence of continuous energy in their spectrographic representations, and although several notes may occur in a given motif, their pattern is usually highly stereotyped between successive renditions of the same motif. One can thus consider starling song as a sequence of motifs, where each motif is an acoustically complex event. The number of unique motifs that a male starling can sing (i.e. his repertoire size) can be quite large, and consequently different song bouts from the same male are not necessarily composed of the same set of motifs. This broad acoustical variation in their song provides several potential cues that starlings might use when learning to recognize the songs of an individual conspecific and while maintaining that recognition over time. One straightforward recognition mechanism is the association of specific motifs with specific singers. Although some sharing of motifs does occur among captive males (Hausberger, 1997; Hausberger & Cousillas, 1995), the motif repertoires of different males living in the wild are generally unique (Adret-Hausberger & Jenkins, 1988; Eens, Pinxten, & Verheyen, 1989; Eens et al., 1991). Thus, learning which males sing which motifs can provide discriminative cues for song recognition. Similar strategies may underlie individual song recognition behaviors observed in other species where individual possess elaborate vocal repertoires such as song sparrows (Beecher, Campbell, & Burt, 1994; Stoddard, Beecher, Horning, & Campbell, 1991), but other strategies are available as well (Brooks & Falls, 1975a, 1975b; Nelson, 1989).

Data from operant studies in starlings indicate that song recognition is based at the level of the motif. Starlings trained using operant conditioning to recognize sets of songs from conspecific individuals can readily generalize correct recognition to novel songs from the same singers (Gentner & Hulse, 1998, 2000b). However, recognition falls to chance when these novel song bouts have no motifs in common with the training songs (Gentner & Hulse, 2000b). This failure to generalize correct recognition to songs composed of novel motifs, or to single novel motifs, is consistent with the notion that starlings learn to recognize the songs of individual conspecific singers by attending to information contained at (or below) the level of the motif. That is, they appear to associate distinct sets of motifs (or variant motif features) with individual singers.

If starlings learn to recognize individuals by the sets of unique motifs that they sing, then once learned, it should be possible to control recognition systematically by varying the proportions of motifs in a given bout that come from two "vocally familiar" males. That is, recognition behavior ought to follow the proportional distribution of motifs from two vocally familiar males rather than the presence or absence of single diagnostic motifs from either male. The behavioral data confirm this prediction. When starlings are compelled to classify conspecific songs, they do so by memorizing large numbers of unique song components (i.e. motifs) and organizing subsets of these motifs into separate classes (Gentner & Hulse, 2000b). Consistent with this idea, recent results demonstrate that motifs form perceptually salient auditory objects embedded in a hierarchy of acoustic patterns (Gentner, 2008). When the sub-motif acoustic structure is permuted in varying ways, recognition of the songs is significantly impaired. These effects are only seen for familiar motifs, however, suggesting that starlings learn the sub-motif structure of motifs and then perceive them as holistic auditory objects (Gentner, 2008).

Additional support for the notion that motifs form holistic auditory objects comes from recent results in starlings that mirror perceptual restoration studies in humans (Seeba & Klump, 2009). In the classic demonstration of perceptual restoration (Warren, 1970), units of speech (phonemes) within a presented sentence are perceived as present by human observers despite being replaced by noise (e.g. a human cough). The restoration effect only occurs when the phoneme is replaced by noise; when silence is inserted instead of noise, phonemes are correctly perceived as missing. This shows that with appropriate contextual cues, observers perceive missing acoustic information at the level of a whole phoneme, suggesting that they have a memory for phonemes as holistic auditory units. Similarly, starlings trained to detect small changes in motif similarities report significantly greater similarity between intact motifs and those with sections replaced by noise when compared to motifs with sections replaced by silence. Importantly, these results hold only for familiar motifs, suggesting that memory for motifs as learned objects informs their perception at later times (Seeba & Klump, 2009). These results provide another line of evidence for considering motifs as acoustic units and suggest that the representation of these motifs emerges via a combination of feed forward auditory input and memory for the motif as an object.

As a cognitive recognition strategy, classifying songs according to their component (motif) structure represents a straightforward method of dealing with these complex acoustic signals. Because individual starlings tend to possess unique motif repertoires, disjoint sets of motifs will generally correspond to individual identity. Therefore, attending to the motif structure captures a significant portion of the individual variation in the signal, albeit at the expense of a heavy demand on memory capacity. From a human perspective, this might seem a suboptimal strategy for individual vocal recognition, because in humans the total set of vocal signals that can be produced by a single individual is theoretically infinite. Instead, individual talker recognition in humans relies on voice characteristics such as timbre, glottal pulsation frequency, and spectral contours imparted by laryngeal morphology (Bricker & Pruzansky, 1976) that are largely independent of the linguistic content (Remez, Fellowes, & Rubin, 1997). Although recent results indicate that some voice characteristics are present in starling song (Gentner, 2008), these birds appear to recognize individuals by preferentially using information at longer temporal scales rather than the short (fast) timescale features that carry voice characteristics. This observation is consistent with the conjecture that the incorporation of individual identity information into the phonetics of the signal may be a plesiomorphic (evolutionarily older) state for vocal communication. As the size of the lexicon and sharing of signal elements between individuals increased, the association of unique vocabularies with different individuals would become increasingly inefficient, forcing the differentiation of vocal identity from other semantic channels.

3.3. Syntactic pattern sensitivity

Part of what makes human language unique is its incorporation of long timescale patterning rules (grammars) that constrain the meaning of specific utterances, e.g. telling us precisely who did what to whom. Animal vocalizations lack the combinatorial grammatical complexity of human language. Yet many animal communication signals have a rich temporal structure, and understanding how the temporal patterns in such signals are used and perceived is of broad interest. Such studies can serve to emphasize important similarities and differences in underlying mechanisms for processing acoustically complex vocal signals.

The prior sections on individual vocal recognition in starlings establish the functional importance of motif-level song organization. Under normal conditions, however, single motifs are almost never produced in isolation but typically occur as part of long and elaborate song bouts where 25–30 different motifs may be strung together in close succession. In the following sections we discuss the role of temporal motif patterning in starling song perception.

3.3.1. Motif sequences
Sensitivity to the temporal patterning of motifs could arise in its simplest form based on the explicit sequences in which familiar motifs are produced. As an example, a bird might learn to recognize motifs A–D (where letters denote different motifs) *and* the sequence ABCD. If this is the case, then presenting the same motifs in a different order may affect song recognition. Consistent with this, starlings trained to recognize naturally patterned songs have a much harder time recognizing novel songs from familiar singers when motifs are strung together in a random order, than when motifs follow each singer's natural motif ordering (Gentner & Hulse, 1998), indicating that starlings do attend to motif sequencing.

3.3.2. Motif patterns
Additional studies of temporal pattern perception in starlings move beyond simple transition probabilities between adjacent motifs to ask if starlings can acquire abstract rules that describe the patterning of familiar motifs (Gentner, Fenn, Margoliash, & Nusbaum, 2006). That is, instead of learning the explicit sequence of motifs ABCD as above, the study asked if starlings could learn the pattern *ABCD*, where the letters now denote *sets* of motifs that can occur at each position. Rules that describe sequences of patterned strings have a rich history in the theory of formal grammars, and form the basis for theories of syntax in human language. Formal grammars can be classified hierarchically according to the complexity of the patterns they can produce or recognize (Hopcroft & Ullman, 1979). Finite-state grammars are among the most limited type of formal grammar, and have been thought to describe all animal communication systems (Hauser, Chomsky, & Fitch, 2002; but see Suzuki, Buck, & Tyack, 2006). Human languages minimally require a grammar more complex than finite-state, called a context-free grammar, in part to support the hierarchical embedding common in many syntactic structures (Chomsky, 1957; Hopcroft & Ullman, 1979).

Starlings can learn to classify patterned strings of song motifs generated by both a finite-state grammar and a context-free grammar (Gentner et al., 2006). In this case, strings generated by the context-free grammar followed the form A^2B^2, while the strings generated by the finite-state grammar followed the form $(AB)^2$, where A and B refer to sets of acoustically distinct starling song

motifs known as "rattles" and "warbles". It is important to note that the context-free grammar learned in this study, while more computationally complex than the finite-state grammar, is not sufficient to fully capture the hierarchical structure in human languages. Specifically, it lacks any long-range dependencies between specific elements that, in for example some English sentences, help define the pairings between subjects and verbs, and thereby enable understanding of who did what to whom. Therefore, one should not conclude from this study that songbirds have the capacity to master "human" grammar. Nonetheless, it is clear that they can extract patterning rules from strings of vocal signals, and that those rules can have a remarkable level of complexity. Whether or not birds and humans rely on similar neuronal mechanisms to extract temporal pattern rules is unknown. In any case, the basic ability to attend to high-level patterns within strings of vocal signals, including at least some patterns defined by formal grammars, does not appear to be a recent or unique adaptation.

Although there has been some discussion of the role of syntactic rules in structuring bird song in chickadees and wrens (Clucas, Freeberg, & Lucas, 2004; Hailman & Ficken, 1986; Holland, Dabelsteen, & Paris, 2000), the existence of complex pattering rules in songbird vocalizations has not received adequate attention. Currently, there is no strong evidence to support the notion that songbirds (or any non-human species) use syntax to vary the *semantic* content of vocal signals in the combinatorial manner. By itself, the ability to process simple syntactic structures may be of little functional significance or adaptive value, and may represent a necessary but insufficient precondition to the use of unbounded signal sets observed in humans. These sorts of primitive pattern recognition abilities might need to be paired with more sophisticated gestural systems as a permissive step to subsequent evolution of human language processing capacities in their modern forms.

4. The neural coding in the songbird auditory system

The behavioral data yield several hypotheses regarding the neural mechanisms underlying the recognition of natural (i.e. complex) acoustic signals. First, the functionality of auditory objects in recognition behavior implies an explicit and discrete representation for vocal signals in the central nervous system. That is, the response functions of single neurons or of populations of neurons in appropriate auditory regions should reflect the segmentation of song at the level of the motif (in starlings) or other behaviorally relevant temporal scales. Second, because vocal recognition behavior requires the *learned* association between songs and singers, the neural representations of motifs should reflect their explicit behavioral relevance. That is, there should be a bias for representations of familiar compared to unfamiliar motifs. Third, the representational mechanisms and capacity (i.e. memory) of the system should permit the acquisition of very large numbers of acoustically complex, natural objects (motifs). The nature of object representation, representational plasticity, and memory capacity are central questions for any researcher interested in the neural coding of natural stimuli, including human speech and language. Below, we consider these questions in the context of the neural representation of starling song.

4.1. The avian auditory system

Auditory information enters the avian brain via sensory transduction in the inner ear, where vibrations on the basilar membrane of the cochlea are transduced into the action potentials propagated to the cochlear nuclei in the brainstem (Fig. 2). The cochlear nuclei are characterized by strong tonotopy, and the neurons within these nuclei have a tendency to fire tonic bursts that align with ampli-

avian auditory system
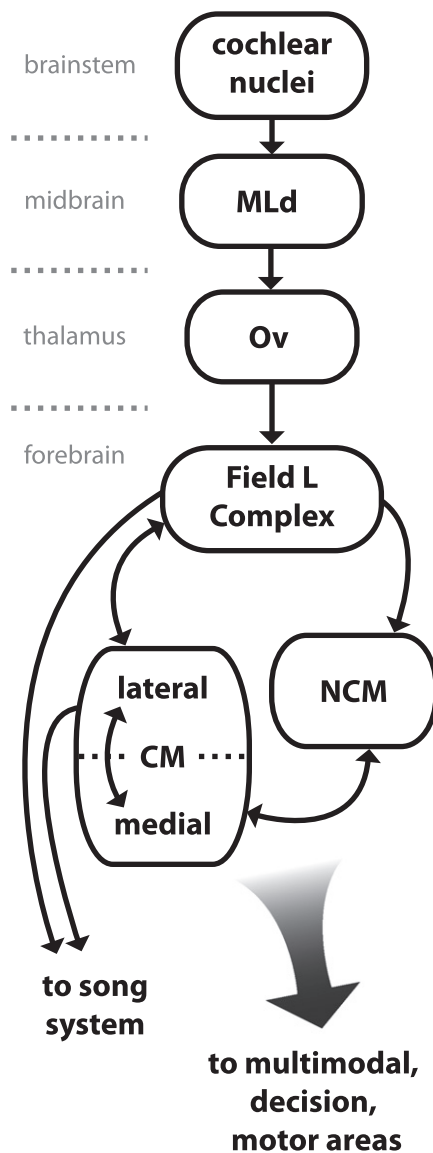
Fig. 2. Schematic of the major auditory processing regions in the avian brain. Auditory information projects from the ear to the brainstem cochlear nuclei. In general, more complex acoustic information is represented in more central structures, i.e. those further from the cochlea. Apart from those to the classic song-control system, projections from the primary auditory pathways to areas responsible for multimodal integration and non-singing behaviors are uncharacterized, but presumably exist. See the text for more details. Thinner arrows indicate direct neural projections. Abbreviations: CM: caudal mesopallium, MLd: mesencephalicus lateralis dorsalis, NCM: caudomedial nidopallium, Ov: nucleus ovoidalis.

tude peaks in the sound pressure waveform of the stimulus (Sachs & Sinnott, 1978). Afferents from the brainstem converge in the midbrain structure mesencephalicus lateralis dorsalis (MLd). MLd is anatomically analogous to the mammalian inferior colliculus (IC), and in many respects it is also functionally analogous. Like the mammalian IC, MLd neurons show a wide variety of tuning characteristics, are not particularly sensitive to level differences, and have temporally precise responses that preserve the timing information in the acoustic signal (Woolley & Casseday, 2004). They also respond to white noise and conspecific song in a way that suggests selectivity for particular spectrotemporal features, and the vast majority of these cells follow sinusoidal amplitude

modulations – suggesting again that these cells encode the temporal features of an acoustic signal with high fidelity (Woolley, Fremouw, Hsu, & Theunissen, 2005).

MLd sends afferents into the avian auditory thalamus, the Nucleus Ovoidalis (Ov). Ov shows a strong tonotopic arrangement, and its neurons have high spontaneous firing rates, and respond to stimuli at their preferred frequency with a tonic increase in firing rate (Bigalke-Kunz, Rübsamen, & Dörrscheidt, 1987). There is also evidence of lateral inhibition in this region, as many cells show suppressive sidebands in their frequency tuning curves (Bigalke-Kunz et al., 1987). Thalamic efferents project from Ov mainly into the Field L complex (shortened here to Field L), the avian analog of the mammalian primary auditory cortex (Karten, 1968, 1969). Specifics of cytoarchitecture and neural connectivity allow further subdivision of the Field L complex into regions L1, L2a, L2b, L3 and L (here 'L' refers to a subregion of the larger complex Fortune & Margoliash, 1992; Saini & Leppelsack, 1977, 1981). L2a is the primary thalamorecipient zone, but L1 and L3 also receive weak thalamic input (Vates, Broome, Mello, & Nottebohm, 1996). Field L shows strong tonotopy and robust responses to both natural song and pure tones (Capsius & Leppelsack, 1996, 1999; Haüsler, 1996; Leppelsack & Vogt, 1976; Sen, Theunissen, & Doupe, 2001).

From Field L, forebrain processing proceeds along a number of parallel and interconnected pathways. Field L itself projects to the caudomedial nidopallium (NCM), which surrounds Field L anatomically, and to the lateral portion of the caudal mesopallium (CLM) (Vates et al., 1996). The medial portion of the caudal mesopallium (CMM) shares reciprocal projections with both CLM and NCM (Vates et al., 1996). Neurons throughout the songbird auditory telencephalon show complex patterns of tonotopic organization (Capsius & Leppelsack, 1996; Haüsler, 1996; Leppelsack & Schwartzkopff, 1972; Rübsamen & Dörrscheidt, 1986), and varying selectivity to species-specific vocalizations (Bonke, Bonke, & Scheich, 1979; Leppelsack & Vogt, 1976; Müller & Scheich, 1985). Neurons in Field L2a/b are less selective for species-specific vocalizations than those in L1 and L3 (Theunissen & Doupe, 1998; Theunissen, Sen, & Doupe, 2000; Theunissen et al., 2004), which in turn are less selective than those in NCM and CM (Grace, Amin, Singh, & Theunissen, 2003; Müller & Scheich, 1985). Increasing song selectivity across these regions is coincident with an increase in the non-linear components of the neural responses (Sen et al., 2001), and higher regions in the pathway (i.e. NCM and CMM) are involved in the representation of complex acoustic features (e.g. Bonke et al., 1979; Gentner, Hulse, & Ball, 2004; Gentner & Margoliash, 2003; Leppelsack & Vogt, 1976; Müller & Scheich, 1985).

Additional support for the role of NCM and CM in the processing of conspecific song comes from studies of stimulus driven expression of the immediate-early-gene (IEG) ZENK, a commonly used marker of song-evoked auditory activity (Mello, Velho, & Pinaud, 2004a) and song-induced experience-dependent plasticity (Jarvis, Mello, & Nottebohm, 1995; Jones et al., 2001; Mello, Nottebohm, & Clayton, 1995; Mello, Velho, & Pinaud, 2004b; Mello, Vicario, & Clayton, 1992; Ribeiro, Cecchi, Magnasco, & Mello, 1998). In starlings, the IEG response in NCM appears tied to stimulus novelty, whereas IEG activity in CMM appears to correlate with the ongoing recognition of familiar songs (Gentner et al., 2004). Together, these observations are consistent with a functional hierarchical organization of the songbird auditory system that is tuned throughout to conspecific song (Hsu, Woolley, Fremouw, & Theunissen, 2004; Woolley, Gill, Fremouw, & Theunissen, 2009; Woolley et al., 2005).

In addition to connections between regions within the auditory system, anatomical (Vates et al., 1996) and functional (Bauer et al., 2008; Shaevitz & Theunissen, 2007) connections have been observed between auditory processing areas and the song production system, providing a path for feedback during the acquisition and maintenance of a bird's own song. While the full range of efferent targets for the auditory telencephalon is unknown, it is likely that brain regions outside the classic "song system" are involved in the subsequent processing of song representations, including areas responsible for integrating multimodal information that naturally accompanies hearing song, as well as regions necessary for carrying out decision processes and motor programs necessary for executing song-mediated behaviors, e.g. copulation displays.

An increasing amount is known about the functionality and gross anatomical connectivity of the structures described above, but there is still little knowledge of the underlying functional micro-circuitry. Early EM and light microscopy studies of local circuitry within Field L indicate that neurons are arranged in small clusters (Saini & Leppelsack, 1977), and it is tempting to hypothesize that these clusters may form functional units. Perhaps related to this, large numbers of GABAergic cells are distributed throughout Field L, NCM, and CMM. In NCM and CMM many of these cells also show song-inducible expression of ZENK (Pinaud et al., 2004), suggesting that inhibitory mechanisms play an important functional role in the representation of learned acoustic material (see the following for further discussions of inhibitory processing in the avian auditory forebrain Capsius & Leppelsack, 1996; Pinaud & Terleph, 2008; Pinaud et al., 2008). Much work remains to be done to test these hypotheses. In general, elucidating the structure and activity of local circuitry will be crucial for understanding how song representation differs in distinct brain regions, the transformations that occur between these regions, and the role of these different representations in a bird's behavior. This will be an area of considerable research effort over the next decade.

The avian forebrain lacks the gross anatomical structure of mammalian cortex. Nonetheless, birds and mammals use their auditory systems to solve many similar behavioral problems. Structural and functional similarities present in the auditory regions of both classes of animals at the cellular (Saini & Leppelsack, 1981) and systems levels (Wild, Karten, & Frost, 1993), (Sockman, Gentner, & Ball, 2002), and across the whole of the forebrain (Jarvis et al., 2005) may reflect these common behavioral goals, Observations suggesting a link between avian studies and the rich body of research on the mammalian cortex highlight the need for comparative studies of auditory system function. Comparing anatomical and functional similarities between birds and mammals and across species within these classes will lend insight into which processes are likely to be fundamental to the representation of the auditory world by a neural substrate, and which may be idiosyncratic to a particular organism or group of organisms.

4.2. Experience-dependent representation of song

The acoustic environment of a songbird is not likely to be constant over time, or between individuals. Territories change hands and the available range of mates fluctuates. In order to effectively interpret the acoustic features in song, the auditory system needs to structure its representation in such a way that allows for dynamic behavioral goals of the organism. A combination of gene expression and electrophysiological studies has helped to characterize several regions within the auditory telencephalon where representations of learned song are thought to be stored.

Many neurons in NCM and CM show a rapid and selective up-regulation of the ZENK (or its protein product) in response to the presentation of conspecific songs (Mello et al., 1992), in manner sensitive to the acoustics of particular song syllables (reviewed in Pinaud & Terleph, 2008; Ribeiro et al., 1998; Terleph, Mello, & Vicario, 2006). Interestingly, these IEG responses (Mello et al., 1995), as well as NCM multiunit recordings (Chew, Vicario, & Nottebohm, 1996), habituate to the repeated presentation of the same conspecific song. Similar habituation in the ZENK response in NCM also

appears closely tied to the context under which song is presented (Kruse, Stripling, & Clayton, 2004). In addition, IEG responses are elevated during specific components of vocal recognition (Gentner et al., 2004), female choice (Gentner, Hulse, Duffy, & Ball, 2001), and other song-based learned behaviors (Jarvis et al., 1995) including vocal acquisition (Phan, Pytte, & Vicario, 2006). While these habituation processes are likely to be important for understanding how the system processes conspecific song (e.g. Dong & Clayton, 2009), non-associative forms of learning are insufficient to capture the complexity of most song-based natural behaviors where reinforcement and stimulus associations are vital.

An organism's sensory experiences evoke long-lasting changes that bias cortical circuits towards representations of behaviorally relevant signals (Gilbert, Sigman, & Crist, 2001). Experience-dependent plasticity in neural responses underlies many forms of associative learning and has been observed across a variety of vertebrate sensory systems and brain regions (e.g. Bakin & Weinberger, 1990; Kay & Laurent, 1999; Kilgard, 2003). Experience dependant plasticity has also been observed in starlings that have been trained, using methods similar to those described above, to recognize sets of conspecific songs. In birds trained this way, single neurons in CMM respond much more robustly (i.e. emit many more action potentials) to songs the birds have learned to recognize than to unfamiliar conspecific songs. The strong response bias for familiar songs is consistent in animals trained with either a two-alternative choice procedure, where all of the training songs are paired with positive reinforcement, or go/no-go procedure, where only half the training songs are paired with positive reinforcement. In the latter case, however, songs associated with positive reinforcement elicited significantly stronger responses from CMM neurons than those associated with no reinforcement (Gentner & Margoliash, 2003). Thus, plasticity in CMM neuronal responses cannot be driven simply by song exposure. Other processes that mark the relevance of each song independent of any particular signal acoustics, such as attention, motivation and/or specific reinforcement, are likely involved.

CMM neurons do not respond at high rates to the entire ensemble of familiar songs. Instead, many individual neurons are selective for only a single song, or for only a small number of motifs within a single song (Gentner & Margoliash, 2003). Thus the same features that drive individual vocal recognition of starling song at the behavioral level (i.e. motifs) also elicit selective responses from single neurons. The mechanisms whereby this selectivity arises over the course of learning are not clear. It may be that the receptive fields of individual neurons are modified to match the spectrotemporal acoustics of behaviorally relevant signals. Rapid shifts in auditory receptive fields driven by behavioral goals have been observed in other systems (Fritz, Elhilali, & Shamma, 2007; Fritz, Shamma, Elhilali, & Klein, 2003). Alternatively, neurons that best match the acoustics on relevant songs may be somehow selected from an otherwise static population, and/or the responses neurons that don't match the relevant acoustics suppressed. These possibilities, along with the larger question of how to best model the receptive fields of these high-level auditory neurons, remain the topic of future research (e.g. Gill, Woolley, Fremouw, & Theunissen, 2008). Nonetheless, recent work consistent with the ideas above suggests that there may be a variety of forms of experience-dependent plasticity throughout the songbird auditory forebrain (Caporello & Gentner, 2008; Jeanne, Sharpee, & Gentner, 2008; Thompson & Gentner, 2008). Spike trains in Field L already carry sufficient information to discriminate between conspecific songs (Narayan, Graña, & Sen, 2006), suggesting that representation of the relevant acoustics may not be the primary role for processing within these regions. Rather, the system may be primarily concerned with the extraction of task-relevant features and their mapping onto appropriate responses.

The observations of experience-dependent plasticity in the songbird auditory system suggest a system that is simultaneously constrained in its immediate representational capacities by each animal's history, yet is tremendously adaptive in its ability to acquire a broad (currently undetermined) range of complex representations. It is tempting to speculate that similar experience-dependent, hierarchical processes may give rise to the phonetic, phonemic, syllabic and word-level representations that must be correlated with language experience in humans. At the same time, one must be careful to temper these kinds of speculations with an appropriate appreciation for the natural history of the organism. While the song recognition system expresses tremendous plasticity, it is reasonable to ask whether all of the links between auditory features and the behaviorally relevant phenomena they represent must be learned during an individual bird's lifetime. The existence of, for example, species-specific warning calls and the preference for conspecific song suggest that the meaning associated with some signals may not require any sort of learning. By extension, such processes may entail substantially different mechanisms at both neural and behavioral levels.

In many songbirds, the females' decision of who to mate with (termed female choice) is driven by the acoustics of competing males' songs. (Catchpole & Slater, 1995; Searcy, 1992). These behaviors provide an interesting context within which to explore the relationship between learned and innate signal processing mechanisms, as they are strongly influenced by both kinds of information. Untrained female starlings show strong preferences for the long songs of older males (Gentner & Hulse, 2000a), and these preferences are reflected in ZENK protein expression levels in NCM (Gentner et al., 2001; see also Maney, MacDougall-Shackleton, MacDougall-Shackleton, Ball, & Hahn, 2003). Yet, the magnitude of these preferences can be modified with relatively little experience (Sockman, Gentner, & Ball, 2005; Sockman et al., 2002). Lesions to CM, the same region involved in song recognition in starlings, lead female zebra finches to produce copulation solicitations to both conspecific and heterospecific song (Macdougall-Shackleton et al., 1998). Similar effects are reported in female canaries with lesions to the song nucleus HVC (Brenowitz, 1991), suggesting that brain regions involved in song production may also function in song perception. Future work should continue to target the neural mechanisms that support this natural decision behavior.

5. Conclusions

Although no two species are identical, this review describes a number of ways in which the auditory processing capacities of songbirds are similar to those of other vertebrates, and in particular to those of humans. Perhaps one reason why birdsong has captivated humans for so many centuries is that we "hear the world" in very similar ways. Indeed, the fact that birds and humans share a similar perceptual representation of the acoustic world holds the promise that higher-level auditory features relevant to more complex behaviors in birds may inhabit a similar acoustic space to those that are relevant for vocal communication in humans. In turn, there may be common neural and behavioral mechanisms adaptive for processing signals within this shared acoustic space. After human language, birdsong is arguably the most acoustically complex and diverse communication signal known. Although it is interesting in it own right, the parallels to human perception make it an excellent model for investigations for higher-level auditory processes, and enable a full range of neurobiological and behavioral techniques that are difficult to use in humans.

Although it is likely that the bird's brain can help us understand important strategies for solving auditory perception tasks that are mutual between humans and birds, birdsong is not language.

In particular, birdsong lacks the rich combinatorial structure and semantics of human vocal signals, and may be more analogous to speech than language (Doupe & Kuhl, 1999). By some theories, much of what makes human language unique is modality independent (e.g. Hauser & Bever, 2008), and so the greatest differences between humans and other animals may lie in the mapping between the outputs of speech processing brain areas and higher-order structures. The mechanisms for extracting, parsing, and representing patterned strings of vocal signals may be conserved across a wide range of species. Within this context, there remain a number of substantial challenges for the songbird auditory system that are germane to our general understanding of systems neuroscience. How are the complete receptive fields for high-level auditory (or other sensory modality) neurons best modeled? The best current models perform poorly beyond the primary thalamorecipient zones. What is the micro-circuitry of Field L, NCM, CM? The organization of the avian auditory telencephalon may be substantially different from that of mammalian cortex, yet very similar function is achieved. How? What are the sources – behavioral, anatomical and physiological – for top-down modulation through attention, reinforcement, memory, etc. within the system? Finally, what are the roles of judgment and decision in the auditory telencephalon and its efferent targets? Further studies into the behavioral and neural basis of the birdsong auditory system promise to be instrumental in answering these questions.

References

Adret-Hausberger, M., & Jenkins, P. F. (1988). Complex organization of the warbling song in starlings. *Behaviour, 107*, 138–156.

Appeltants, D., Gentner, T. Q., Hulse, S. H., Balthazart, J., & Ball, G. F. (2005). The effect of auditory distractors on song discrimination in male canaries (*Serinus canaria*). *Behavioural Processes, 69*(3), 331–341.

Bakin, J. S., & Weinberger, N. M. (1990). Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the guinea pig. *Brain Research, 536*(1–2), 271–286.

Bauer, E. E., Coleman, M. J., Roberts, T. F., Roy, A., Prather, J. F., & Mooney, R. (2008). A synaptic basis for auditory-vocal integration in the songbird. *Journal of Neuroscience, 28*(6), 1509–1522.

Bee, M., & Klump, G. M. (2004). Primitive auditory stream segregation: A neurophysiological study in the songbird forebrain. *Journal of Neurophysiology, 92*(2), 1088–1104.

Beecher, M., Campbell, S. E., & Burt, J. (1994). Song perception in the song sparrow: Birds classify by song type but not by singer. *Animal Behaviour, 47*, 1343–1351.

Bigalke-Kunz, B., Rübsamen, R., & Dörrscheidt, G. J. (1987). Tonotopic organization and functional characterization of the auditory thalamus in a songbird, the European starling. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 161*(2), 255–265.

Bonke, B. A., Bonke, D., & Scheich, H. (1979). Connectivity of the auditory forebrain nuclei in the guinea fowl (*Numida meleagris*). *Cell and Tissue Research, 200*(1), 101–121.

Bregman, A. S. (1990). *The auditory scene auditory science analysis: The perceptual organization of sound*. Cambridge: MIT Press.

Brenowitz, E. A. (1991). Altered perception of species-specific song by female birds after lesions of a forebrain nucleus. *Science, 251*(4991), 303–305.

Bricker, P. D., & Pruzansky, S. (1976). Speaker recognition. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 295–326). New York: Academic Press.

Brooks, R. J., & Falls, J. B. (1975a). Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbors and strangers. *Canadian Journal of Zoology, 53*, 879–888.

Brooks, R. J., & Falls, J. B. (1975b). Individual recognition by song in white-throated sparrows. III. Song features used in individual recognition. *Canadian Journal of Zoology, 53*, 1749–1761.

Brown, C. H., & Maloney, C. G. (1986). Temporal integration in two species of old world monkeys: Blue monkeys (*Cercopithecus mitis*) and grey-cheeked mangabeys (*Cercocebus albigena*). *The Journal of the Acoustical Society of America, 79*(4), 1058–1064.

Caporello, E., Gentner, T. Q. (2008). *Song recognition learning drives experience-dependent changes in the avian telencephalic auditory area field L*. Paper presented at the society for neuroscience.

Capsius, B., & Leppelsack, H. J. (1996). Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. *Hearing Research, 96*(1–2), 59–70.

Capsius, B., & Leppelsack, H. (1999). Response patterns and their relationship to frequency analysis in auditory forebrain centers of a songbird. *Hearing Research, 136*(1–2), 91–99.

Catchpole, C. K., & Slater, P. B. J. (1995). *Bird song*. Cambridge: Cambridge University Press.

Chew, S. J., Vicario, D. S., & Nottebohm, F. (1996). A large-capacity memory system that recognizes the calls and songs of conspecifics. *Proceedings of the National Academy of Science, 93*, 1950–1955.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Clucas, B. A., Freeberg, T. M., & Lucas, J. R. (2004). Chick-a-dee call syntax, social context, and season affect vocal responses of Carolina chickadees (*Poecile carolinensis*). *Behavioral Ecology and Sociobiology, 57*, 187–196.

Cynx, J., & Nottebohm, F. (1992). Testosterone facilitates some conspecific song discriminations in castrated zebra finches (*Taeniopygia guttata*). *Proceedings of the National Academy of Sciences of the United States of America, 89*(4), 1376–1378.

Demany, L., & Semal, C. (1989). Detection thresholds for sinusoidal frequency modulation. *Journal of the Acoustical Society of America, 85*(3), 1295–1301.

Dong, S., & Clayton, D. F. (2009). Habituation in songbirds. *Neurobiology of Learning and Memory, 92*(2), 183–188.

Dooling, R. J. (1980). Behavior and psychophysics of hearing in birds. In A. N. Popper & R. R. Fay (Eds.), *Comparative studies of hearing in vertebrates*. New York: Springer-Verlag.

Dooling, R. J. (1982). *Auditory perception in birds. Acoustic communication in birds* (Vol. 1). Academic Press.

Dooling, R. J. (1992). Hearing in birds. In D. Webster, R. Fay, & A. N. Popper (Eds.), *The Evolutionary Biology of Hearing* (pp. 545–559). New York: Springer-Verlag.

Dooling, R. J. (2000). Hearing in birds and reptiles. In R. J. Dooling, R. R. Fay, & A. N. Popper (Eds.), *Comparative hearing: Birds and reptiles* (pp. 308–359). New York: Springer-Verlag.

Dooling, R. J., Brown, S. D., Klump, G. M., & Okanoya, K. (1992). Auditory perception of conspecific and heterospecific vocalizations in birds: Evidence for special processes. *Journal of Comparative Psychology, 106*(1), 20–28.

Dooling, R. J., Okanoya, K., Downing, J., & Hulse, S. H. (1986). Hearing in the starling (*Sturnus vulgaris*): Absolute thresholds and critical ratios. *Bulletin of the Psychonomic Society, 24*(6), 462–464.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22*, 567–631.

Eens, M., Pinxten, M., & Verheyen, R. F. (1989). Temporal and sequential organization of song bouts in the European starling. *Ardea, 77*, 75–86.

Eens, M., Pinxten, M., & Verheyen, R. F. (1991). Male song as a cue for mate choice in the European starling *Sturnus vulgaris*. *Behaviour, 116*(3–4), 210–238.

Fastl, H. (1978). Frequency discrimination for pulsed versus modulated tones. *Journal of the Acoustical Society of America, 63*(1), 275–277.

Fay, R. R. (1988). *Hearing in vertebrates: A psychophysics databook*. Winnetka: Hill – Fay Associates.

Fortune, E. S., & Margoliash, D. (1992). Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (*Taenopygia guttata*). *Journal of Comparative Neurology, 325*(3), 388–404.

Fritz, J. B., Elhilali, M., & Shamma, S. A. (2007). Adaptive changes in cortical receptive fields induced by attention to complex sounds. *Journal of Neurophysiology, 98*(4), 2337–2346.

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience, 6*(11), 1216–1223.

Gentner, T. Q. (2008). Temporal scales of auditory objects underlying birdsong vocal recognition. *Journal of the Acoustical Society of America, 124*(2), 1350–1359.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204–1207.

Gentner, T. Q., & Hulse, S. H. (1998). Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*. *Animal Behaviour, 56*(3), 579–594.

Gentner, T. Q., & Hulse, S. H. (2000a). Female European starling preference and choice for variation in conspecific male song. *Animal Behaviour, 59*(2), 443–458.

Gentner, T. Q., & Hulse, S. H. (2000b). Perceptual classification based on the component structure of song in European starlings. *Journal of the Acoustical Society of America, 107*(6), 3369–3381.

Gentner, T. Q., Hulse, S. H., & Ball, G. F. (2004). Functional differences in forebrain auditory regions during learned vocal recognition in songbirds. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 190*(12), 1001–1010.

Gentner, T. Q., Hulse, S. H., Duffy, D., & Ball, G. F. (2001). Response biases in auditory forebrain regions of female songbirds following exposure to sexually relevant variation in male song. *Journal of Neurobiology, 46*(1), 48–58.

Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature, 424*(6949), 669–674.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron, 31*(5), 681–697.

Gill, P., Woolley, S. M. N., Fremouw, T. E., & Theunissen, F. E. (2008). What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. *Journal of Neurophysiology, 99*(6), 2809–2820.

Grace, J. A., Amin, N., Singh, N. C., & Theunissen, F. E. (2003). Selectivity for conspecific song in the zebra finch auditory forebrain. *Journal of Neurophysiology, 89*(1), 472–487.

Hailman, J. P., & Ficken, M. S. (1986). Combinational animal communication with computable syntax: Chick-a-dee calling qualifies as 'language' by structural linguistics. *Animal Behaviour, 34*, 1899–1901.

Hausberger, M. (1997). Social influences on song acquisition and sharing in the European starling (*Sturnus vulgaris*). In C. Snowden & M. Hausberger (Eds.),

Social influences on vocal development (pp. 128–156). Cambridge: Cambridge University Press.

Hausberger, M., & Cousillas, H. (1995). Categorization in birdsong: From behavioural to neuronal responses. *Behavioural Processes, 35*(1–3), 83–91.

Hauser, M. D., & Bever, T. (2008). Behavior: A biolinguistic agenda. *Science, 322*(5904), 1057–1059.

Haüsler, U. (Ed.). (1996). *Measurement of short-time spatial activity patterns during auditory stimulation in the starling.* New York, NY: Plenum Press.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579.

Holland, J., Dabelsteen, T., & Paris, A. L. (2000). Coding in the song of the wren: Importance of rhythmicity, syntax and element structure. *Animal Behaviour, 60*(4), 463–470.

Hopcroft, J., & Ullman, J. (1979). *Introduction to automata theory, languages, and computation.* Reading, MA: Addison-Wesley.

Hsu, A., Woolley, S. M., Fremouw, T. E., & Theunissen, F. E. (2004). Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. *Journal of Neuroscience, 24*(41), 9201–9211.

Hulse, S. H., MacDougall-Shackleton, S. A., & Wisniewski, A. B. (1997). Auditory scene analysis by songbirds: Stream segregation of birdsong by European starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology, 111*(1), 3–13.

Jarvis, E. D., Gunturkun, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience, 6*(2), 151–159.

Jarvis, E. D., Mello, C. V., & Nottebohm, F. (1995). Associative learning and stimulus novelty influence the song-induced expression of an immediate early gene in the canary forebrain. *Learning and Memory, 2*, 62–80.

Jeanne, J. M., Sharpee, T. O., & Gentner, T. Q. (2008). *Song recognition learning drives experience-dependent representations in the auditory forebrain region CLM.* Paper presented at the society for neuroscience.

Jones, M. W., Errington, M. L., French, P. J., Fine, A., Bliss, T. V., Garel, S., et al. (2001). A requirement for the immediate early gene Zif268 in the expression of late LTP and long-term memories. *Nature Neuroscience, 4*(3), 289–296.

Karten, H. J. (1968). The ascending auditory pathway in the pigeon (*Columba livia*). II. Telencephalic projections of the nucleus ovoidalis thalami. *Brain Research, 11*, 134–153.

Karten, H. J. (1969). The organization of the avian telencephalon and some speculations on the phylogeny of the amniote telencephalon. In J. M. Petras & C. R. Noback (Eds.). *Comparative and evolutionary aspects of the vertebrate central nervous system* (Vol. 167, pp. 164–179). New York: The New York Academy of Science.

Kay, L. M., & Laurent, G. (1999). Odor- and context-dependent modulation of mitral cell activity in behaving rats. *Nature Neuroscience, 2*(11), 1003–1009.

Kilgard, M. (2003). Cholinergic modulation of skill learning and plasticity. *Neuron, 38*(5), 678–680.

Klump, G. M., & Langemann, U. (1995). Comodulation masking release in a songbird. *Hearing Research, 87*(1–2), 157–164.

Klump, G. M., & Maier, E. H. (1989). Gap detection in the starling (*Sturnus vulgaris*). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 164*(4), 531–538.

Kroodsma, D. E., & Miller, E. H. (Eds.). (1996). *Ecology and evolution of acoustic communication in birds.* Ithaca: Comstock/Cornell.

Kruse, A. A., Stripling, R., & Clayton, D. F. (2004). Context-specific habituation of the ZENK gene response to song in adult zebra finches. *Neurobiology of Learning and Memory, 82*(2), 99–108.

Langemann, U., & Klump, G. M. (1992). Frequency discrimination in the European starling (*Sturnus vulgaris*): A comparison of different measures. *Hearing Research, 63*(1–2), 43–51.

Langemann, U., Klump, G. M., & Dooling, R. J. (1995). Critical bands and critical-ratio bandwidth in the European starling. *Hearing Research, 84*(1–2), 167–176.

Leppelsack, H. J., & Vogt, M. (1976). Responses of auditory neurons in the forebrain of a songbird to stimulation with species-specific sounds. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 107*(3), 263–274.

Leppelsack, H.-J., & Schwartzkopff, J. (1972). Properties of acoustic neurons in the caudal neostriatum of birds. *Journal of Comparative Physiology, 80*, 137–140.

MacDougall-Shackleton, S. A., Hulse, S. H., Gentner, T. Q., & White, W. (1998). Auditory scene analysis by European starlings (*Sturnus vulgaris*): Perceptual segregation of tone sequences. *Journal of the Acoustical Society of America, 103*(6), 3581–3587.

Maier, E. H., & Klump, G. M. (1990). Auditory duration discrimination in the European starling (*Sturnus vulgaris*). *Journal of the Acoustical Society of America, 88*(2), 616–621.

Maney, D. L., MacDougall-Shackleton, E. A., MacDougall-Shackleton, S. A., Ball, G. F., & Hahn, T. P. (2003). Immediate early gene response to hearing song correlates with receptive behavior and depends on dialect in a female songbird. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 189*(9), 667–674.

Mello, C., Nottebohm, F., & Clayton, D. (1995). Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon. *Journal of Neuroscience, 15*, 6919–6925.

Mello, C. V., Velho, T. A., & Pinaud, R. (2004a). Song-induced gene expression: A window on song auditory processing and perception. *Annals of the New York Academy of Sciences, 1016*, 263–281.

Mello, C. V., Velho, T. A. F., & Pinaud, R. (2004b). Song-induced gene expression: A window on song auditory processing and perception. *ANYAS, 1016*, 263–281.

Mello, C. V., Vicario, D. S., & Clayton, D. F. (1992). Song presentation induces gene-expression in the songbird forebrain. *Proceedings of the National Academy of Sciences of the United States of America, 89*(15), 6818–6822.

Müller, S. C., & Scheich, H. (1985). Functional organization of the avian auditory field L. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 156*(1), 1–12.

Narayan, R., Best, V., Ozmeral, E., McClaine, E., Dent, M., Shinn-Cunningham, B., et al. (2007). Cortical interference effects in the cocktail party problem. *Nature Neuroscience, 10*(12), 1601–1607.

Narayan, R., Graña, G., & Sen, K. (2006). Distinct time scales in cortical discrimination of natural sounds in songbirds. *Journal of Neurophysiology, 96*(1), 252–258.

Nelson, D. A. (1988). Feature weighting in species song recognition by the field sparrow (*Spizella pusilla*). *Behaviour, 106*(1–2), 158–181.

Nelson, D. A. (1989). Song frequency as a cue for recognition of species and individuals in the field sparrow (*Spizella pusilla*). *Journal of Comparative Psychology, 103*(2), 171–176.

Nelson, D. A., & Marler, P. (1990). The perception of birdsong and an ecological concept of signal space. In W. C. Stebbins & M. A. Berkely (Eds.). *Comparative perception: Complex signals* (Vol. 2, pp. 443–477). New York: John Wiley and Sons.

Okanoya, K., & Dooling, R. J. (1987). Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds. *Journal of Comparative Psychology, 101*(1), 7–15.

Phan, M. L., Pytte, C. L., & Vicario, D. S. (2006). Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proceedings of the National Academy of Sciences of the United States of America, 103*(4), 1088–1093.

Pinaud, R., & Terleph, T. A. (2008). A songbird forebrain area potentially involved in auditory discrimination and memory formation. *Journal of Biosciences, 33*(1), 145–155.

Pinaud, R., Terleph, T. A., Tremere, L. A., Phan, M. L., Dagostin, A. A., Leão, R. M., et al. (2008). Inhibitory network interactions shape the auditory processing of natural communication signals in the songbird auditory forebrain. *Journal of Neurophysiology, 100*(1), 441–455.

Pinaud, R., Velho, T. A. F., Jeong, J. K., Tremere, L. A., Leão, R. M., von Gersdorff, H., et al. (2004). GABAergic neurons participate in the brain's response to birdsong auditory stimulation. *European Journal of Neuroscience, 20*(5), 1318–1330.

Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology – Human Perception and Performance, 23*(3), 651–666.

Ribeiro, S., Cecchi, G. A., Magnasco, M. O., & Mello, C. V. (1998). Toward a song code: Evidence for a syllabic representation in the canary brain. *Neuron, 21*(2), 359–371.

Rübsamen, R., & Dörrscheidt, G. J. (1986). Tonotopic organization of the auditory forebrain in a songbird, the European starling. *Journal of Comparative Physiology, 158*, 639–646.

Sachs, M. B., & Sinnott, J. M. (1978). Responses to tones of single cells in nucleus magnocellularis and nucleus angularis of the redwing blackbird (*Agelaius phoeniceus*). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 126*(4), 347–361.

Saini, K., & Leppelsack, H. (1977). Neuronal arrangement in the auditory field L of the neostriatum of the starling. *Cell and Tissue Research, 176*(3), 309–316.

Saini, K. D., & Leppelsack, H. J. (1981). Cell types of the auditory caudomedial neostriatum of the starling. *Journal of Comparative Neurology, 198*, 209–230.

Searcy, W. A. (1992). Song repertoire and mate choice in birds. *American Zoologist, 32*, 71–80.

Seeba, F., & Klump, G. M. (2009). Stimulus familiarity affects perceptual restoration in the European starling (*Sturnus vulgaris*). *PLoS One, 4*(6), e5974.

Sen, K., Theunissen, F. E., & Doupe, A. J. (2001). Feature analysis of natural sounds in the songbird auditory forebrain. *Journal of Neurophysiology, 86*(3), 1445–1458.

Shaevitz, S. S., & Theunissen, F. E. (2007). Functional connectivity between auditory areas field L and CLM and song system nucleus HVC in anesthetized zebra finches. *Journal of Neurophysiology, 98*(5), 2747–2764.

Sinnott, J. M., Sachs, M. B., & Hienz, R. D. (1980). Aspects of frequency discrimination in passerine birds and pigeons. *Journal of Comparative and Physiological Psychology, 94*(3), 401–415.

Sockman, K. W., Gentner, T. Q., & Ball, G. F. (2002). Recent experience modulates forebrain gene-expression in response to mate-choice cues in European starlings. *Proceedings of the Royal Society of London Series B – Biological Sciences, 269*(1508), 2479–2485.

Sockman, K. W., Gentner, T. Q., & Ball, G. F. (2005). Complementary neural systems for the experience-dependent integration of mate-choice cues in European starlings. *Journal of Neurobiology, 62*(1), 72–81.

Stoddard, P. K. (1996). Vocal recognition of neighbors by territorial passerines. In D. E. Kroodsma & E. H. Miller (Eds.), *Ecology and evolution of acoustic communication in birds* (pp. 356–374). Ithaca: Comstock/Cornell.

Stoddard, P. K., Beecher, M. D., Horning, C. L., & Campbell, S. E. (1991). Recognition of individual neighbors by song in the song sparrow, a species with song repertoires. *Behavioral Ecology and Sociobiology, 29*, 211–215.

Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (2000). Call-note discriminations in black-capped chickadees (*Poecile atricapillus*). *Journal of Comparative Psychology, 114*(4), 357–364.

Suzuki, R., Buck, J. R., & Tyack, P. L. (2006). Information entropy of humpback whale songs. *JASA, 119*(3), 1849–1866.

Terleph, T. A., Mello, C. V., & Vicario, D. S. (2006). Auditory topography and temporal response dynamics of canary caudal telencephalon. *Journal of Neurobiology, 66*(3), 281–292.

Theunissen, F. E., Amin, N., Shaevitz, S. S., Woolley, S. M., Fremouw, T., & Hauber, M. E. (2004). Song selectivity in the song system and in the auditory forebrain. *Annals of the New York Academy of Sciences, 1016*, 222–245.

Theunissen, F. E., & Doupe, A. J. (1998). Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVc of male zebra finches. *Journal of Neuroscience, 18*(10), 3786–3802.

Theunissen, F. E., Sen, K., & Doupe, A. J. (2000). Spectral–temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *Journal of Neuroscience, 20*(6), 2315–2331.

Thompson, J. V., & Gentner, T. Q. (2008). *Song recognition learning elicits differential suppression of neural responses to learned songs in NCM*. Paper presented at the society for neuroscience.

Vates, G. E., Broome, B. M., Mello, C. V., & Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. *Journal of Comparative Neurology, 366*(4), 613–642.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*(3917), 392–393.

Wild, J. M., Karten, H. J., & Frost, B. J. (1993). Connections of the auditory forebrain in the pigeon (*Columba livia*). *Journal of Comparative Neurology, 337*(1), 32–62.

Woolley, S. M. N., & Casseday, J. H. (2004). Response properties of single neurons in the zebra finch auditory midbrain: Response patterns, frequency coding, intensity coding, and spike latencies. *Journal of Neurophysiology, 91*(1), 136–151.

Woolley, S. M. N., Fremouw, T. E., Hsu, A., & Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature Neuroscience, 8*(10), 1371–1379.

Woolley, S. M., Gill, P. R., Fremouw, T., & Theunissen, F. E. (2009). Functional groups in the avian auditory system. *Journal of Neuroscience, 29*(9), 2780–2793.